

A mixed PDE-Monte Carlo approach for pricing credit default index swaptions

Vlad Bally¹, Lucia Caramellino^{2*}, Antonino Zanette³

¹ Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne-la-Vallée
e-mail: bally@univ-mlv.fr

² Dipartimento di Matematica, Università di Roma *Tor Vergata*
e-mail: caramell@mat.uniroma2.it

³ Dipartimento di Finanza dell'Impresa e dei Mercati Finanziari, Università di Udine
e-mail: antonino.zanette@uniud.it

Received: 19 October 2005 / Accepted: 29 May 2006 – © 2006 The Authors¹

Abstract. The problem of numerically pricing credit default index swaptions on a large number of names is considered. We place ourselves in a stochastic intensity framework, where Ornstein-Uhlenbeck-type correlated processes are used to model both firms' distance to default and a macroeconomic state variable. Here the default of the firms' follows the reduced-form approach and the (random) intensity of the default depends on the behavior of the diffusion processes. We propose here a numerical method based on both a Monte Carlo and a deterministic approach for solving PDEs by finite differences. Numerical tests demonstrate the efficiency and the robustness of the proposed procedure.

Journal of Economic Literature Classification: D81, C40

Mathematics Subject Classification (2000): 91B28, 60H30, 65C05

1. Introduction

The credit risk market has been growing rapidly in recent years. In particular, credit default swaps (CDSs) are one of the most actively traded credit derivatives (see, e.g., Duffie (1999), Duffie, Saita and Wang (2006), Schönbucher (2003)).

*Corresponding author

¹This is a “Springer Open Choice” article. Unrestricted non-commercial use, distribution, and reproduction in any medium is permitted, provided the original author and source are credited.

A CDS provides protection in the event of default (called a credit event) of a specific company (called the reference entity). It is then an agreement between two counterparties, allowing the first to be “long” on a third-party credit risk and the second to be “short” on the credit risk. More precisely, one has first to introduce a reference asset, which is typically a credit-risky bond issued by a third-party corporation. Now, the two counterparties, say A and B, enter into an agreement: A pays to B a fixed periodic coupon for the specified life of the reference asset. Party B makes no payments unless the default event occurs for the reference asset. If such a credit event occurs, B makes a payment to party A, and the swap then terminates. The size of the payment is usually linked to the decline in the reference asset’s market value following the credit event.

In this paper, we consider the problem of pricing options on a portfolio of a large number of names (typically 100), where the underlying instrument for each name is a CDS. Such risky financial instruments are known in the literature as *index default swaptions*, or *credit default index swaptions*, or also *CDS index swaptions*; the last is the denomination we use (see, e.g., Jackson (2005)).

A CDS index swaption is an option to buy or sell the underlying CDSs at a specified date. A payer swaption gives the holder of the option the right to buy protection (pay premium) and a receiver swaption gives the holder of the option the right to sell protection (receive premium). This basic definition of a CDS index swaption is similar to that of a CDS option, an option on a single-entity CDS, but a CDS index swaption is significantly different from a CDS option. In the case of an option on a single-entity, if the reference entity defaults before the option’s expiry, the option is knocked out and becomes worthless. For a CDS index swaption, when a reference entity defaults before the option’s expiry, the loss is paid by the protection seller to the protection buyer when the option is exercised. Even if there is only one entity in the portfolio, a CDS index swaption is still different from a single-entity CDS option: if the entity defaults before the option’s expiry, the option’s seller pays to the protection buyer the lost amount at expiry. Clearly, a CDS index swaption is always more valuable than a single-entity CDS option.

In this paper, we allow the underlying possible defaults to follow an intensity-based model, in which the default time is a stopping time with a given intensity process (as in Duffie and Singleton (1999) or also Schönbucher (2003); different approaches can be found in literature; see, e.g., Jackson (2005)). In our application, the default intensities depend on firms’ distance to default and on observable variables linked with the likelihood of default, such as macroeconomic variables related to the business cycle. In our framework, both firms’ distance to default and the macroeconomic

variable are modelled through correlated Ornstein-Uhlenbeck processes, with suitable coefficients. The model, which is described mathematically below (Section 2), was first introduced in Duffie and Wang (2004) and later generalized by Duffie, Saita and Wang (2006).

Our main contribution is the construction of an algorithm for the pricing of CDS index swaptions on a large number of names. It is based on a mixing between classical numerical approximation by partial differential equations (PDEs) and Monte Carlo approaches. For the model considered here, this algorithm offers an efficient alternative to a pure vanilla Monte Carlo method, which is the classical tool to handle such derivatives but which becomes unfeasible from a computational point of view when the number of the underlying firms is large. Comparisons with results from a pure vanilla Monte Carlo method are presented here mainly in order to benchmark the results of our algorithm. It would be interesting to set up other pure and improved Monte Carlo techniques, such as control variates, which might lead to better computational efficiency. But, to our knowledge, no results are available in this direction, although this could be an interesting research problem.

The advantage of our algorithm is that it circumvents the nested simulation problem of a standard Monte Carlo algorithm by approximating conditional expectations using a PDE approach. In fact, the PDE approach is particularly suited to problems which are controlled by two-dimensional diffusions, and can be efficiently handled by standard numerical techniques, such as finite differences. We emphasize the fact that the model we use, and in particular the modelled correlations among the underlying stochastic processes, allows us to reduce to a number of different two-dimensional problems equal to the number of the firms considered.

Lastly, even if the main goal of our work is to introduce a new methodology (the mixed PDE-Monte Carlo algorithm) constructed ad hoc for the pricing of a complex credit product (CDS index swaption), nevertheless we hope that this contribution helps to pave the way for future work addressing more challenging credit risk problems in higher dimension in a dynamical setting.

The paper is organized as follows. In Section 2 we introduce the model of the stochastic intensity process for the default with a description of CDSs and CDS index swaptions. Section 3 presents the mixed PDE-Monte Carlo approach for pricing CDS index swaptions. Numerical results are presented in Section 4.

2. The model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbb{P})$ be a filtered probability space, denoting the “risk neutral world”, where the following (diffusion) processes are defined:

– $(Y_t)_{t \in [0, T]}$, modelled as

$$dY_t = \kappa_Y(\theta_Y - Y_t)dt + \sigma_Y dW_t^Y; \quad (1)$$

– as $i = 1, \dots, n$, $(D_{i,t})_{t \in [0, T]}$, modelled as

$$dD_{i,t} = \kappa_D(\theta_{D,i} - D_{i,t})dt + \sigma_D(\rho dW_t^c + \sqrt{1 - \rho^2} dW_t^i). \quad (2)$$

In (1) and (2), the processes $W^Y, W^c, W^1, \dots, W^n$ denote independent Brownian motions; ρ stands for a correlation coefficient, thus belonging to $(-1, 1)$; $\kappa_Y, \theta_Y, \sigma_Y$ and $\kappa_D, \theta_{D,i}, \sigma_D$ are all constant. Therefore, Y and D_1, \dots, D_n are Ornstein-Uhlenbeck-type (correlated) diffusions. Notice that Y is independent of D_1, \dots, D_n , while D_1, \dots, D_n depend on each other and such a dependence is given by the correlation ρ .

The filtration is then $\mathcal{F}_t = \sigma((W_s^Y, W_s^c, W_s^1, \dots, W_s^n); s \leq t)$.

The intuitive meaning of the process Y arises from the fact that it models a macroeconomic state variable, e.g., the U.S. personal income growth, while D_1, \dots, D_n model the firms' distance to default (for details, see Duffie and Wang (2004) or Duffie, Saita and Wang (2006)). Recall that the distances to default are correlated: in fact, the common Brownian motion W^c , which appears in all the SDEs giving D_1, \dots, D_n (see (2)), captures the correlation among the individual distances to default.

We assume that the time of default follows the intensity-based approach: with τ_i denoting the default instant for the i th firm, then its (random) failure intensity λ^i depends on both Y and D_i as

$$\lambda_t^i = \Lambda(Y_t, D_{i,t}), \quad \text{where } \Lambda(y, d) = \exp(\mu_0 + \mu_1 y + \mu_2 d). \quad (3)$$

Note that the constant parameters μ_0, μ_1 and μ_2 are common to all firms.

Recall that the above definitions mean that, for $s < t$,

$$\mathbb{P}(\tau_i > t \mid \mathcal{F}_s) = \mathbb{E}(e^{-\int_s^t \Lambda(Y_u, D_{i,u}) du} \mid \mathcal{F}_s) \quad \text{on the set } \{\tau_i > s\},$$

where, from now on, the symbol \mathbb{E} denotes the expectation under \mathbb{P} .

For the construction of τ_i , we refer to Duffie and Singleton (1999), Schönbucher (2003) and also Lando (1999). We only recall here that this defines the default time as the first jump instant of a Cox process with intensity process λ^i , that is, $\tau_i = \inf\{t \geq 0 : \int_0^t \lambda_s^i ds = Z\}$, where Z is an exponential random variable, of parameter 1, independent of all the Brownian motions already defined.

Remark 1. We make some remarks about our choice for the parameter modelling. Indeed, concerning the distance-to-default processes, we assume a common mean reversion κ_D and a common volatility σ_D (see (2)); moreover, the intensity failures are defined through a function with common parameters, i.e., μ_0, μ_1 and μ_2 (see (3)). Such a choice might appear very restrictive,

but this parsimonious model tries to overcome the problem of an extremely high-dimensional state-vector, consisting in one macroeconomic covariate, personal income growth Y , and the distance to default D_i for each firm i among the n in total. On the other hand, a relevant parameter concerning the distance to default, that is, the long-run mean parameter $\theta_{D,i}$ (see (2)), varies firm by firm. Lastly, the assumed homogeneity of correlation across different firms makes the model numerically tractable.

We now describe a CDS index swaption mathematically. This is an option to buy protection on the CDS index at rate K , which pays at time t

$$\max \left(\sum_{i=1}^n V_{i,t}, 0 \right),$$

where $V_{i,t}$ is the market value at time t of a default swap on name i at rate K (notice that here K does not depend on i). Recall that a CDS is a credit derivative that protects its owner from the default event of the issuer of the underlying CDS.

Thus, if r denotes the risk-free interest rate, the market value at time 0 of such an option is given by

$$P_0 = e^{-rt} \mathbb{E} \left(\max \left(\sum_{i=1}^n V_{i,t}, 0 \right) \right). \quad (4)$$

The values of the quantities $V_{1,t}, \dots, V_{n,t}$ depend on the fact that the default has or has not been observed at time t as follows:

- if $\tau_i \leq t$, that is, name i has defaulted by time t , then $V_{i,t}$ is the loss L_i in the case of the default of name i :

$$V_{i,t} = L_i;$$

- if $\tau_i > t$, that is, no default occurred up to t , then

$$V_{i,t} = B_{i,t} - A_{i,t},$$

where $B_{i,t}$ is the value at time t of the future payments by the seller of protection at default and $A_{i,t}$ is the value at time t of the future payments of the buyer of protection.

To give an explicit expression for $B_{i,t}$ and $A_{i,t}$, we introduce further notation. Set $t = t_0 < t_1 < \dots < t_N = T$, the premium payment dates, which are such that $t_{j+1} - t_j = \Delta t$ for any j (usually, Δt is equal to three months). Define

$$p_{i,j,t} = \mathbb{P}(\tau_i > t_j | \mathcal{F}_t) = \mathbb{E} \left(\exp \left(- \int_t^{t_j} \Lambda(Y_s, D_{i,s}) ds \right) \middle| \mathcal{F}_t \right)$$

as the probability of survival from t to t_j given \mathcal{F}_t on $\{\tau_i > t\}$. Finally, set

$$d_j = \exp(-r(t_j - t))$$

as the time discount factor from time t to time t_j . Then

$$\begin{aligned} B_{i,t} &= \sum_{j=1}^N d_j (p_{i,j-1,t} - p_{i,j,t}) \mathbb{E}(L_i), \\ A_{i,t} &= \sum_{j=1}^N d_j p_{i,j,t} K. \end{aligned} \tag{5}$$

To resume, for any i , the market value $V_{i,t}$ at time t of a credit default swap on name i at rate K is given by

$$V_{i,t} = L_i \mathbf{1}_{\{\tau_i \leq t\}} + (B_{i,t} - A_{i,t}) \mathbf{1}_{\{\tau_i > t\}},$$

where $B_{i,t}$ and $A_{i,t}$ are defined in (5).

Remark 2. In principle, the loss L_i can be modelled in many different ways. Here, L_i is a random variable distributed between zero and one according to a specified law, which is the uniform one. More elaborate laws could be used in practice, but we wish to benchmark a simple case. One can allow L_i to be correlated with other variables but, for the sake of simplicity, we assume here that L_i is independent of all else.

3. The mixed PDE-Monte Carlo numerical approach

Consider a derivative whose price P_0 is given by (4). To numerically compute P_0 , consider a Monte Carlo approach, that is,

$$P_0 = e^{-rt} \mathbb{E} \left(\max \left(\sum_{i=1}^n V_{i,t}, 0 \right) \right) \simeq \frac{e^{-rt}}{M} \sum_{m=1}^M \max \left(\sum_{i=1}^n V_{i,t}^{(m)}, 0 \right),$$

where the index m stands for one of the M trials and $V_{i,t}^{(m)}$ denotes the m th simulation for $V_{i,t}$. Then the problem is how to simulate

$$V_{i,t} = L_i \mathbf{1}_{\{\tau_i \leq t\}} + (B_{i,t} - A_{i,t}) \mathbf{1}_{\{\tau_i > t\}}.$$

Obviously, the r.v. $V_{i,t}$ cannot be exactly replicated, so initially we have to discuss an approximation $\tilde{V}_{i,t}$ of $V_{i,t}$.

3.1. Checking if default does or does not occur

First of all, we have to check if the default of the i th firm does or does not occur up to time t . In order to do this, we have to use the default probabilities, which of course depend on the paths followed by the process Y and D_i on $[0, t]$. Therefore, one has first to discretize the time interval $[0, t]$ in order to set up an Euler scheme for the two processes and consequently to compute the default probabilities. To be more precise, set $0 = s_0 < s_1 < \dots < s_\ell = t$ a discretization of $[0, t]$ such that $s_{k+1} - s_k = \Delta s$ is “small enough”. Define \bar{Y} and $\bar{D}_{i,t}$ the (discrete) Euler schemes for Y and D_i respectively by:

$$\begin{aligned}\bar{Y}_0 &= Y_0, \quad \bar{D}_{i,0} = D_{i,0} \quad \text{and, for } k = 1, \dots, \ell, \\ \bar{Y}_k &= \bar{Y}_{k-1} + \kappa_Y(\theta_Y - \bar{Y}_{k-1})\Delta s + \sigma_Y \Delta W_k^Y, \\ \bar{D}_{i,k} &= \bar{D}_{i,k-1} + \kappa_D(\theta_{D,i} - D_{i,k-1})\Delta s + \sigma_D(\rho \Delta W_k^c + \sqrt{1 - \rho^2} \Delta W_k^i).\end{aligned}$$

Here, ΔW_k^Y , ΔW_k^c and ΔW_k^i denote the variation between s_{k-1} and s_k of the Brownian motions W^Y , W^c and W^i respectively. We stress that, in our notation, $\bar{Y}_k \simeq Y_{s_k}$ and $\bar{D}_{i,k} \simeq D_{i,s_k}$. Once the above schemes are set, we can approximate the event “the default has occurred up to time t ” as follows. First, observe that the probability of no default of the firm i up to s_k given that the firm is still alive at time s_{k-1} can be approximated by using the approximation (\bar{Y}, \bar{D}_i) as follows: on the set $\{\tau_i > s_{k-1}\}$ one has

$$\mathbb{P}(\tau_i > s_k \mid \mathcal{F}_{s_{k-1}}) = \mathbb{E}\left(e^{-\int_{s_{k-1}}^{s_k} \Lambda(Y_s, D_{i,s})ds} \mid \mathcal{F}_{s_{k-1}}\right) \simeq e^{-\Lambda(\bar{Y}_{k-1}, \bar{D}_{i,k-1})\Delta s}.$$

Then, if we set

$$\bar{q}_{i,k} = e^{-\Lambda(\bar{Y}_{k-1}, \bar{D}_{i,k-1})\Delta s},$$

- one first computes $\bar{q}_{i,0}$ and: with probability $\bar{q}_{i,0}$ one says that no default occurred up to s_1 ; with probability $1 - \bar{q}_{i,0}$ one says that default did occur up to s_1 ;
- if the default did not occur up to s_{k-1} , then one computes $\bar{q}_{i,k}$ and: with probability $\bar{q}_{i,k}$ one says that no default occurred up to s_k ; with probability $1 - \bar{q}_{i,k}$ one says that default did occur up to s_k .

Recall that in practice this means generating r.v.’s $Z_{i,k}$, uniformly distributed on $(0, 1)$, independent of each other and of any other random variable written above, and to check if $Z_{i,k}$ is or is not less than $\bar{q}_{i,k}$: if $Z_{i,k} < \bar{q}_{i,k}$ then no default is assumed to have occurred; if $Z_{i,k} > \bar{q}_{i,k}$ then the default is assumed to have occurred.

This procedure leads to an approximation $\bar{\mathbf{1}}_i$ of the random variable $\mathbf{1}_{\{\tau_i > t\}}$: we set $\bar{\mathbf{1}}_i = 1$ if no default has been observed up to $s_\ell = t$, otherwise $\bar{\mathbf{1}}_i = 0$. Once $\bar{\mathbf{1}}_i$ is obtained, we set:

- $V_{i,t} \simeq \bar{V}_{i,t} = L_i$ if $\bar{1}_i = 0$;
- $V_{i,t} \simeq \bar{V}_{i,t} = (B - A)_{i,t}$ if $\bar{1}_i = 1$,

where $\overline{(B - A)}_{i,t}$ is a suitable approximation for $B_{i,t} - A_{i,t}$, to be discussed in Sections 3.2 and 3.3.

3.2. Standard Monte Carlo approach if default does not occur

In this section, we assume that we have observed no default up to t , that is, it is on the set $\{\tau_i > t\}$. Then, we have to approximate $B_{i,t} - A_{i,t}$. By formula (4), we can rewrite

$$\begin{aligned} B_{i,t} - A_{i,t} &= \sum_{j=1}^N d_j (p_{i,j-1,t} - p_{i,j,t}) \mathbb{E}(L_i) - \sum_{j=1}^N d_j p_{i,j,t} K \\ &= \sum_{j=1}^N c_{i,j} p_{i,j,t} + c_{i,0}, \end{aligned} \quad (6)$$

where

$$c_{i,j} = \begin{cases} \mathbb{E}(L_i) d_1 & \text{if } j = 0, \\ -(K + \mathbb{E}(L_i)) d_j + \mathbb{E}(L_i) d_{j+1} & \text{if } 1 \leq j \leq N-1, \\ -(K + \mathbb{E}(L_i)) d_N & \text{if } j = N \end{cases} \quad (7)$$

(we take $p_{i,0,t} = 1$).

This means that $B_{i,t} - A_{i,t}$ is a linear combination of the functions

$$p_{i,j,t} = \mathbb{E} \left(\exp \left(- \int_t^{t_j} \Lambda(Y_s, D_{i,s}) ds \right) \middle| \mathcal{F}_t \right).$$

Now, since the pair $(X_{i,s})_s = (Y_s, D_{i,s})_s$ is a diffusion, then the Markov property holds, so that

$$p_{i,j,t} = P_{i,j}(t, X_{i,t}), \quad \text{where} \quad P_{i,j}(t, x) = \mathbb{E}_{t,x} \left(\exp \left(- \int_t^{t_j} \Lambda(X_{i,s}) ds \right) \right), \quad (8)$$

where as usual $\mathbb{E}_{t,x}$ denotes expectation among the paths starting at x at time t . Therefore, $p_{i,j,t}$ has a representation in terms of an expectation, so that one can try to compute it by the standard Monte Carlo method: letting $\bar{X}_{i,t} = (\bar{Y}_t, \bar{D}_{i,t})$ be the approximation already observed for $X_{i,t}$, then

$$p_{i,j,t} \simeq P_{i,j}(t, \bar{X}_{i,t}) \simeq \frac{1}{M'} \sum_{m'=1}^{M'} \exp \left(- \int_t^{t_j} \Lambda(\bar{X}_{i,s}^{(m')}) ds \right), \quad (9)$$

where $(\bar{X}_{i,s}^{(m')})_{s \geq t}$ denotes a simulation of an approximation $(\bar{X}_{i,s})_{s \geq t}$ (e.g., by an Euler scheme) for $(X_{i,s})_{s \geq t}$ on the time interval $[t, T]$ starting from $\bar{X}_{i,t}$ at time t (and obviously, the integral is approximated by Riemann sums).

Therefore, summing up, we have for any firm i :

- $m = 1, \dots, M$ simulations of the path on $[0, t]$, telling us if default does or does not occur, and if not, giving us a position, say $\bar{X}_{i,t}^{(m)}$, of the pair (Y, D_i) at time t ;
- for any $m = 1, \dots, M$ giving no default, we have to start with M' further simulations of the pair $X_i = (Y, D_i)$ on $[t, T]$ starting at $\bar{X}_{i,t}^{(m)}$, in order to approximate the probabilities as in (9).

This can give rise to $M \times M'$ simulations, for any $i = 1, \dots, n$, so the total number can be as large as $M \times M' \times n$. In other words, an algorithm of this kind is too expensive from a computational point of view. One could try to use other variance reduction techniques, for example, a control variate. But this approach would not overcome the problem of the further M' simulations, leading us to look for an alternative method.

3.3. PDE approach if default does not occur

The idea starts from the fact that deterministic methods to approximate expectations by numerical solutions of PDEs, work quite well if the dimension is small.

We return to $B_{i,t} - A_{i,t}$ and rewrite it as

$$B_{i,t} - A_{i,t} = \varphi_i(t, X_{i,t}), \quad (10)$$

where (see (6)–(8))

$$\varphi_i(t, x) = \sum_{j=1}^N c_{i,j} P_{i,j}(t, x) + c_{i,0}. \quad (11)$$

Now, we propose to compute $\varphi_i(t, X_{i,t})$ numerically as the solution of a parabolic PDE on \mathbb{R}^2 , which in turn is evaluated by finite difference methods. To this purpose, using the Feynman-Kac formula one can first compute $P_{i,j}(t, X_{i,t})$ as follows.

Proposition 1. *One has $P_{i,j}(t, X_{i,t}) = v_{i,j}(t, X'_{i,t})$, where $X'_{i,t} = (Y'_t, D'_{i,t})$ is the two-dimensional diffusion solving the SDE:*

$$\begin{aligned} dY'_t &= -\kappa_Y Y'_t dt + \sigma_Y dW_t^Y, & Y'_0 &= Y_0, \\ dD'_{i,t} &= -\kappa_D D'_{i,t} dt + \sigma_D(\rho dW_t^c + \sqrt{1 - \rho^2} dW_t^i), & D'_{i,0} &= D_{i,0}, \end{aligned} \quad (12)$$

and where $v_{i,j}$ solves

$$\begin{aligned} \partial_t v_{i,j}(s, x) + \mathcal{L}v_{i,j}(s, x) - g_i(s, x)v_{i,j}(s, x) &= 0, & (s, x) &\in (0, t_j) \times \mathbb{R}^2 \\ v_{i,j}(t_j, x) &= 1, \end{aligned} \quad (13)$$

where (recall that $x = (y, d)$)

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\sigma_Y^2\partial_{yy}^2 + \frac{1}{2}\sigma_D^2\partial_{dd}^2 - \kappa_Y y\partial_y - \kappa_D d\partial_d, \\ g_i(s, x) &= q_i(s) \exp(\mu_1 y + \mu_2 d), \quad \text{with} \\ q_i(s) &= \exp\left(\mu_0 + \mu_1\theta_Y(1 - e^{-s\kappa_Y}) + \mu_2\theta_{D,i}(1 - e^{-s\kappa_D})\right).\end{aligned}$$

Remark 3. There is a simple relation connecting the processes $X_i = (Y, D_i)$ and $X'_i = (Y', D'_i)$: for any s ,

$$\begin{aligned}Y'_s &= Y_s - \theta_Y(1 - e^{-\kappa_Y s}) \quad \text{and} \\ D'_{i,s} &= D_{i,s} - \theta_{D,i}(1 - e^{-\kappa_D s}).\end{aligned}$$

The proof is straightforward (but also the next proof as well).

Proof of Proposition 1. It is well-known that the pair $(X_{i,s})_s = ((Y_s, D_{i,s}))_s$ has an explicit expression given by

$$\begin{aligned}Y_s &= Y_0 e^{-\kappa_Y s} + \theta_Y(1 - e^{-\kappa_Y s}) + \sigma_Y \int_0^s e^{-\kappa_Y(u-s)} dW_u^Y, \\ D_{i,s} &= D_{i,0} e^{-\kappa_D s} + \theta_{D,i}(1 - e^{-\kappa_D s}) \\ &\quad + \sigma_D \int_0^s e^{-\kappa_D(u-s)} d(\rho W_u^c + \sqrt{1 - \rho^2} dW_u^i).\end{aligned}$$

Then, with $Y'_s = Y_s - \theta_Y(1 - e^{-\kappa_Y s})$ and $D'_{i,s} = D_{i,s} - \theta_{D,i}(1 - e^{-\kappa_D s})$, it is straightforward to see that the pair $(X'_{i,s})_s = ((Y'_s, D'_{i,s}))_s$ solves (12). Now, for any s ,

$$\begin{aligned}\Lambda(Y_s, D_{i,s}) &= \exp\left(\mu_0 + \mu_1 Y_s + \mu_2 D_{i,s}\right) = \exp\left(\mu_1 Y'_s + \mu_2 D'_{i,s}\right) \cdot q_i(s), \\ \text{where } q_i(s) &= \exp\left(\mu_0 + \mu_1\theta_Y(1 - e^{-s\kappa_Y}) + \mu_2\theta_{D,i}(1 - e^{-s\kappa_D})\right).\end{aligned}$$

Then, for $x = (y, d)$, if $g_i(s, x) = q_i(s) \exp(\mu_1 y + \mu_2 d)$, one has $\Lambda(X_{i,s}) = g_i(s, X'_{i,s})$ and, therefore,

$$\mathbb{E}_{t, X_{i,t}}\left(\exp\left(-\int_t^{t_j} \Lambda(X_{i,s}) ds\right)\right) = \mathbb{E}_{t, X'_{i,t}}\left(\exp\left(-\int_t^{t_j} g_i(s, X'_{i,s}) ds\right)\right).$$

In other words, $P_{i,j}(t, X_{i,t}) = v_{i,j}(t, X'_{i,t})$, where

$$v_{i,j}(t, x) = \mathbb{E}_{t,x}\left(\exp\left(-\int_t^{t_j} g_i(s, X'_{i,s}) ds\right)\right).$$

Now, by the Feynman-Kac formula, it immediately follows that $v_{i,j}$ solves the PDE problem (13). \square

Using Proposition 1, one can represent the function φ_i in (10) and (11) as the solution of a PDE problem, as follows.

Proposition 2. Consider the following backward set of functions $\psi_{i,N}(s, x)$, \dots , $\psi_{i,1}(s, x)$ (where \mathcal{L} and g_i are as defined in Proposition 1):

– $\psi_{i,N}(s, x)$ solves:

$$\begin{aligned} \partial_t \psi_{i,N}(s, x) + \mathcal{L} \psi_{i,N}(s, x) \\ - g_i(s, x) \psi_{i,N}(s, x) = 0, \quad (s, x) \in (0, t_N) \times \mathbb{R}^2, \\ \psi_{i,N}(t_N, x) = c_{i,N}; \end{aligned}$$

– as $j = N - 1, \dots, 1$, $\psi_{i,j}(s, x)$ solves:

$$\begin{aligned} \partial_t \psi_{i,j}(s, x) + \mathcal{L} \psi_{i,j}(s, x) \\ - g_i(s, x) \psi_{i,j}(s, x) = 0, \quad (s, x) \in (0, t_j) \times \mathbb{R}^2, \\ \psi_{i,j}(t_j, x) = c_{i,j} + \psi_{i,j+1}(t_j, x). \end{aligned}$$

Then, $\varphi_i(t, X_{i,t}) = \psi_{i,1}(t, X'_{i,t}) + \mathbb{E}(L_i)d_1$, where X' is defined through (12).

Proof. Consider $\psi_{i,N}$ first: it solves the same PDE as in (13) with $j = N$, with constant (final) Cauchy condition equal to $c_{i,N}$. Then

$$\psi_{i,N}(s, x) = c_{i,N} v_{i,N}(s, x), \quad s \leq t_N.$$

Now, for $j = N - 1, \dots, 1$, again $\psi_{i,j}$ solves the same PDE as in (13), with (final) Cauchy condition equal to $c_{i,j} + \psi_{i,j+1}(t_j, x)$. It then follows that

$$\psi_{i,j}(s, x) = c_{i,j} v_{i,j}(s, x) + \psi_{i,j+1}(s, x), \quad s \leq t_j.$$

Therefore, by iteration,

$$\psi_{i,1}(s, x) = \sum_{j=1}^N c_{i,j} v_{i,j}(s, x), \quad s \leq t_1,$$

so that as $s = t (< t_1)$ one has

$$\psi_{i,1}(t, X'_{i,t}) = \sum_{j=1}^N c_{i,j} v_{i,j}(t, X'_{i,t}) = \sum_{j=1}^N c_{i,j} P_{i,j}(t, X_{i,t}),$$

where the last equality is proved in Proposition 1. Therefore, by (11), one has

$$\varphi_i(t, X_t) = \psi_{i,1}(t, X'_{i,t}) + c_{i,0} = \psi_{i,1}(t, X'_{i,t}) + \mathbb{E}(L_i)d_1,$$

and the statement holds. \square

Remark 4. Observe that one could also avoid considering the new process X'_i and use the original process X_i directly. In such a case, one would obtain the same results, but of course the operator $\partial_t + \mathcal{L} - g_i$ has to be replaced by $\partial_t + \mathcal{L}_i - \Lambda$. It has to be stressed that the advantage of our representation is that the differential operator \mathcal{L} is the same for any i and only g_i changes as i varies. This is important from a computational point of view since this allows us to reduce the number of numerical operations (e.g., in an explicit finite difference scheme).

The result in Proposition 2 is used in practice in the following way. We first compute an approximation $\bar{\varphi}_i$ of φ_i by using deterministic numerical methods for solutions of PDEs (here, for a fixed (firm) i , the associated PDE is on \mathbb{R}^2 , so deterministic methods work efficiently). Therefore, we can finally set the approximation

$$B_{i,t} - A_{i,t} \simeq \overline{(B - A)_{i,t}} = \bar{\varphi}_i(t, \bar{X}_{i,t}).$$

The algorithm can be summarized as follows:

- a) fix the firm i and the m th simulation ($i = 1, \dots, n$ and $m = 1, \dots, M$);
- b) following Section 3.1, check if default does or does not occur on $[0, t]$, and set the m th simulation $\bar{V}_{i,t}^{(m)}$ of $V_{i,t}$ as:
 - b1) if default is observed, then $\bar{V}_{i,t}^{(m)} = L_i$,
 - b2) if no default is observed, then $\bar{V}_{i,t}^{(m)} = \bar{\varphi}_i(t, \bar{X}_{i,t}^{(m)})$;
- c) compute the price by averaging:

$$P_0 \simeq \bar{P}_0 = \frac{e^{-rt}}{M} \sum_{m=1}^M \max \left(\sum_{i=1}^n \bar{V}_{i,t}^{(m)}, 0 \right).$$

It is worth stressing some remarks concerning item b2) above.

We solve the PDEs in Proposition 2 numerically for each name i with a finite difference method. The numerical procedure consists of the following five steps.

1. The set of linear parabolic problems in Proposition 2 have to be localized to a bounded domain in space, that is, \mathbb{R}^2 is in practice replaced by $\Omega_l = (y'_{\min}, y'_{\max}) \times (d'_{\min}, d'_{\max})$. The choice of Ω_l must serve two main purposes. First, Ω_l must be large enough to ensure the convergence of the approximating function to the true one. Second, the values of $(\bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})_m$ (obtained from $(\bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})_m$ by using the transformation as in Remark 3) have to be contained in Ω_l . Then, for each name i , we first simulate all the paths and then construct the grid. Afterwards, we solve the PDEs with a finite difference method.

2. The approximating solutions $\bar{\psi}_{i,N}, \dots, \bar{\psi}_{i,1}$ are computed by means of finite difference methods involving discrete functions. For our numerical studies, we use an explicit finite difference method (see Wilmott, Dewynne and Howison (1993), Villeneuve and Zanette (2001) and we suitably add, as is usually done, Neumann homogenous artificial boundary conditions.
3. We construct for this purpose a time-space grid. Each time interval $(t_{j-1}, t_j]$ is split in subintervals and we consider a mesh of the space domain $[y'_{\min}, y'_{\max}] \times [d'_{\min}, d'_{\max}]$ consisting of subintervals (y'_k, y'_{k+1}) and (d'_k, d'_{k+1}) , where $y'_{\min} = y'_0 < \dots < y'_{M_Y} = y'_{\max}$ and $d'_{\min} = d'_0 < \dots < d'_{M_D} = d'_{\max}$.
4. Since we need $\bar{\varphi}_i$ at time t , we solve the PDEs in Proposition 2 backwards up to t , so that we obtain (see Remark 3)

$$\bar{\varphi}_i(t, y_{j_1}, d_{j_2}) = \bar{\psi}_{i,1}(t, y'_{j_1}, d'_{j_2}) + \mathbb{E}(L_i)d_1, \quad \text{with} \\ y'_{j_1} = y_{j_1} - \theta_Y(1 - e^{-\kappa_Y t}) \quad \text{and} \quad d'_{j_2} = d_{j_2} - \theta_{D,i}(1 - e^{-\kappa_D t}).$$

5. Now, in order to estimate $\bar{\varphi}_i(t, \bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})$, we look for the four points of the space grid which are neighborhood points for $(\bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})$ and then we approximate $\bar{\varphi}_i(t, \bar{Y}_t^{(m)}, \bar{D}_{i,t}^{(m)})$ using bilinear interpolation of $\bar{\psi}_{i,1}$ in these points.

Remark 5. Recall that our use of PDE techniques has been done in order to overcome the problem of the further nested M' simulations for each firm i . This problem could be handled also by using regression in place of PDEs, as was been done for numerically pricing American options (see Carriere (2001) and Longstaff and Schwartz (2001)), a topic which is worth investigating. However, the use of regression for approximating conditional expectations is a good idea for American option pricing purposes, where the time interval is split into small time subintervals. In fact, this ensures that the conditional expectations are made by following the law of a process in a small time interval, and this makes the regression approximation work well. Our context is actually different (see, e.g., (8)), so that a method involving regression is not trivial and must be analyzed in detail.

4. Numerical results

In this section we illustrate numerically the efficiency of the PDE-Monte Carlo approach introduced in Section 3.3 and we compare, for benchmark purposes, this approach with the vanilla Monte Carlo approach of Section 3.2. All the computations were performed in double precision on a PC Pentium 1.73 GHz.

We treat the problem of pricing CDS index swaptions in the model introduced in Section 2 and we consider the following values of the parameters:

- varying number of firms: $n = 50, 100, 125$;
- option parameters: times $t = 1, T = 6$, then $t_j = t + 0.25j$, $j = 1, \dots, 20$; L_i uniformly distributed on $(0, 1)$; interest rate $r = 0.05$; rate $K = 0.001$;
- default probability parameters: $\mu_0 = -4.2017$, $\mu_1 = -0.4597$, $\mu_2 = -0.4411$

As regards the diffusion parameters, we consider two test cases, named A and B :

- **case A:** $\kappa_Y = 0.6524$, $\theta_Y = 1.8901$, $\sigma_Y = 0.8888$, $Y_0 = 3$; $\kappa_D = 0.1185$, $\theta_{D,i} = 4.32 + i/n$, $\sigma_D = 0.9657$, $D_{i,0} = 3$; correlation coefficient $\rho = 0.2684$;
- **case B:** $\kappa_Y = 0.1137$, $\theta_Y = 0.1076$, $\sigma_Y = 1$, $Y_0 = 3$; $\kappa_D = 0.0355$, $\theta_{D,i} = 4.32 + i/n$, $\sigma_D = 0.346$, $D_{i,0} = 3$; correlation coefficient $\rho = 0.2684$;

The choice of these parameters was inspired by a similar combination in Duffie and Wang (2004).

We compare the results obtained by the mixed PDE-Monte Carlo approach and the (standard) Monte Carlo algorithm, in which the following settings are taken into account.

- (i) For the mixed PDE-Monte Carlo approach, we take the number M of simulations to be 10000, the number l of discretization steps, in which the time interval $[0, t]$ is split, to be 32. The PDEs are numerically solved through an explicit finite difference method, with time discretization step set equal to 40, while the spatial discretization step is obtained via the stability condition. We add Neumann homogenous artificial boundary conditions.
- (ii) For the Monte Carlo algorithm, we take a number M of simulations of the first part equal to the number M' of simulations of the second part, both equal to 1000 or 10000; moreover, we split the time interval $[0, t]$ in $\ell = 32$ or 64 subintervals.

The results for the price and the estimated standard deviation, divided by the square root of the number of simulations, for cases A and B are presented in next tables, in the *Price* and *St.Dev./ \sqrt{M}* columns respectively, according to the mixed PDE-Monte Carlo method summarized in (i) (in the tables, PDE-MC) or the vanilla Monte Carlo method as in (ii) (in the table, MC 1000 \times 32 and MC 10000 \times 64). Lastly, reports are given in Tables 1, 2 and 3, for a number n of firms equal to 50, 100 and 125 respectively.

We take the price from the MC 10000 \times 64 case as a benchmark value. In such a case, we observe two different performances for our PDE-MC

Table 1. Number of firms $n = 50$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo method (MC).

| CASE | A | | B | |
|----------------------|-------|---------------------|-------|---------------------|
| | Price | St.Dev./ \sqrt{M} | Price | St.Dev./ \sqrt{M} |
| PDE-MC | 0.038 | 0.002 | 0.094 | 0.003 |
| MC 1000×32 | 0.036 | 0.005 | 0.082 | 0.010 |
| MC 10000×64 | 0.036 | 0.001 | 0.095 | 0.003 |

Table 2. Number of firms $n = 100$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo method (MC).

| CASE | A | | B | |
|----------------------|-------|---------------------|-------|---------------------|
| | Price | St.Dev./ \sqrt{M} | Price | St.Dev./ \sqrt{M} |
| PDE-MC | 0.029 | 0.001 | 0.145 | 0.004 |
| MC 1000×32 | 0.032 | 0.006 | 0.144 | 0.010 |
| MC 10000×64 | 0.024 | 0.001 | 0.141 | 0.005 |

Table 3. Number of firms $n = 125$: price and standard deviation for the mixed PDE-Monte Carlo method (PDE-MC) and the full Monte Carlo method (MC).

| CASE | A | | B | |
|----------------------|-------|---------------------|-------|---------------------|
| | Price | St.Dev./ \sqrt{M} | Price | St.Dev./ \sqrt{M} |
| PDE-MC | 0.024 | 0.001 | 0.166 | 0.006 |
| MC 1000×32 | 0.021 | 0.006 | 0.162 | 0.020 |
| MC 10000×64 | 0.021 | 0.001 | 0.170 | 0.007 |

method. In fact, case B works well, because the relative error is between 1% and 2.8%. Despite this, case A works less well, giving an error higher than that from case B. Nevertheless, it has to be remarked that in case A prices are much smaller, and it is well-known that low values might be responsible for high relative errors.

We make some final remarks on the computational cost in time of the PDE-MC algorithm. The time computation results for $n = 50, 100, 125$ are presented in Table 4. Obviously, comparisons with the time needed by the full Monte Carlo algorithms are not significant (to give an idea, the range is between of hours and days, and in fact the use of the vanilla Monte Carlo approach was made in order to benchmark our results). However, Table 4 shows very low CPU times spent by our algorithm. We can then conclude

Table 4. Time computation for the mixed PDE-Monte Carlo method (PDE-MC).

| | | | |
|-----------------|----------|-----------|-----------|
| number of firms | $n = 50$ | $n = 100$ | $n = 125$ |
| PDE-MC CPU time | 10 sec. | 19 sec. | 23 sec. |

that the proposed mixed PDE-Monte Carlo method can be used to provide a fast and reasonably robust pricing of CDS index swaptions.

5. Conclusions

In this paper, we investigated a new numerical method, a mixed PDE-Monte Carlo algorithm, to approximate the price of credit default index swaptions on a large number of names in the presence of stochastic intensity.

The straightforward application of a pure Monte Carlo algorithm is too expensive from a computational of view. On the contrary, the main appeal of the proposed method is that it allows us to obtain price results for this credit derivatives product cheaply in terms of computational time. The numerical results confirm the reliability of the method.

Based on these results, the mixed PDE-Monte Carlo algorithm presented in this work may be considered to tackle our credit derivative pricing problems in a dynamical setting.

Acknowledgements. The authors are grateful to Darrell Duffie for having suggested the problem and for his encouraging help. The authors also thank the editor Marco Li Calzi and two anonymous referees for their useful comments and suggestions.

References

Carriere, J. (1996): Valuing of the early-exercise price for options using simulations and non-parametric regression. *Insurance: Mathematics & Economics* **19**, 19–30

Duffie, D. (1999): Credit swap valuation. *Financial Analysts Journal* **55**, 73–87

Duffie, D., Wang, K. (2004): Multi-period corporate failure prediction with stochastic co-variates. Preprint.

Duffie, D., Saita, L., Wang, K. (2006): Multi-period corporate failure prediction with stochastic covariates. *Journal of Financial Economics*, to appear.

Duffie, D., Singleton, K. (1999): Modelling term structures of defaultable bonds. *The Review of Financial Studies* **12**, 687–720

Jackson, A. (2005): A new method for pricing index default swaptions. Working paper. Global Credit Derivatives Research, Citigroup, London

Lando, D. (1999): On Cox processes and credit risky securities. *Review of Derivatives Research* **2**, 99–120

- Longstaff, F., Schwartz, E. (2001): Valuing American options by simulations: a simple least-squares approach. *The Review of Financial Studies* **14**, 113–148
- Schönbucher, P.J. (2003): Credit derivatives pricing models: models, pricing, implementation. Wiley, Chichester
- Wilmott, P., Dewynne, J., Howison, S. (1993): Option pricing: mathematical models and computation. Oxford Financial Press, Oxford
- Villeneuve, S., Zanette, A. (2001): Comparison of finite difference methods for pricing American options on two stocks. Working paper 1–2001. Dipartimento di Finanza dell’Impresa e dei Mercati Finanziari, Università di Udine, Udine